DOCUMENT RESUME

ED 189 574 CS 005 555

AUTHOR Weinshank, Annette B.
TITLE An Observational Study of the Relationship between Diagnosis and Remediation in Reading. Research Series No. 72.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Apr 80
CONTRACT 400-76-0073
NOTE 26p.
AVAILABLE FROM Institute for Research on Teaching, College of Education, Michigan State University, 252 Erickson Hall, East Lansing, MI 48824 ($2.75)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis: *Diagnostic Teaching: Elementary Education: Reading Consultants: *Reading Diagnosis: *Reading Research: *Remedial Reading: Teaching Styles

ABSTRACT

A study was conducted to determine whether reading specialists' diagnoses lead directly to remedial recommendations. Eight experienced reading specialists were observed during 24 sessions as they performed a series of tasks using three simulated cases of reading difficulty, two of which were thinly disguised versions of the same reading problem. The study demonstrated that (1) the bulk of the diagnostic and remedial statements for a given case were made only once: (2) across all cases a relatively small number of categories of diagnostic and remedial statements accounted for all statements made more than once: (3) examination of common case information led to neither common diagnoses, common remediations, nor common associations between remediation and diagnosis: (4) only by aggregating diagnostic and remedial statements across clinicians could the outlines of meager consensus on each case be demonstrated: (5) at the individual clinician level, there was essentially no correlation between diagnosis and remediation: (6) at the group level, diagnosis and remediation showed a modest level of association: and (7) clinicians did not follow their stated plans regarding information collection procedures and the writing of the diagnosis and remediation. (Author/AEA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED189574

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Research Series No. 72

AN OBSERVATIONAL STUDY OF THE RELATIONSHIP BETWEEN DIAGNOSIS AND REMEDIATION IN READING

Annette B. Weinshank

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824

April, 1980

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)

CS805555

2

AUG 7 1980

Institute for Research on Teaching

Teachers' thoughts and decisions are the focus of studies currently under way at Michigan State University's Institute for Research on Teaching (IRT). The IRT was founded in April 1976 with a $3.6 million grant from the National Institute of Education. That grant has since been renewed, extending IRT's work through September 1981. Funding is also received from other agencies and foundations. The Institute has major projects investigating teacher decision-making, including studies of reading diagnosis and remediation, classroom management strategies, instruction in the areas of language arts, reading, and mathematics, teacher education, teacher planning, effects of external pressures on teachers' decisions, socio-cultural factors, and teachers' perceptions of student affect. Researchers from many different disciplines cooperate in IRT research. In addition, public school teachers work at IRT as half-time collaborators in research, helping to design and plan studies, collect data, and analyze results. The Institute publishes research reports, conference proceedings, occasional papers, and a free quarterly newsletter for practitioners. For more information or to be placed on the IRT mailing list please write to : The IRT Editor, 252 Erickson, MSU, East Lansing, Michigan 48824.

Director: Judith E. Lanier

Associate Directors: Lawrence W. Lezotte and Andrew C. Porter

Editorial Staff:

Lawrence W. Lezotte, coordinator of Communications/Dissemination
Linda Shalaway, IRT editor
Janet Flegg, assistant editor

(Lee S. Shulman, co-director with Judith E. Lanier from 1976-1978, and director in 1979, is on a one-year leave at the Center for Advanced Study in the Behavioral Sciences, Stanford, California.)

AN OBSERVATIONAL STUDY OF THE RELATIONSHIP BETWEEN DIAGNOSIS AND REMEDIATION IN READING[1]

Annette B. Weinshank[2]

Since early in this century, the overwhelming assertion of the literature on diagnosis and remediation of reading difficulties has been that diagnosis is an indispensible prerequisite to remediation. At present, training procedures for reading specialists continue to stress the necessity of gathering information from a variety of sources, including information from schools, parents, and agencies working with the family; information on student health and developmental history; and information obtained through analytic testing in reading and other areas, including personality. Considerable time and money are spent in reading clinics and schools for information gathering, testing, diagnosis, and remediation. The major justification for all this activity seems to be the assumption that effective remediation can proceed only after a thorough diagnosis has been prepared. What if this assumption is incorrect?

If present diagnosis and remediation procedures are shown to be unrelated, serious issues would be raised about the usefulness of present training programs in the diagnosis and remediation of reading difficulties. The value of extended diagnostic testing sessions and data gathering in

---

[1]This paper is a summary of Annette Weinshank's doctoral dissertation. It was submitted to Michigan State University in 1980.

[2]Annette B. Weinshank is a teacher co-investigator with IRT's Clinical Studies Project and a former research intern with that project. She is also an experienced reading specialist now working in the Lansing, Michigan public schools.

clinics would come under question. The role of the school reading consultant would have to be reevaluated. In sum, demonstrating the inaccuracy of the assumption would provoke marked consequences throughout the educational system.

This study empirically investigated the clinical problem-solving behavior of reading specialists as they examined information contained in simulated cases of reading difficulty, to find out if any reliable relationship existed between their diagnoses and remediations. In order to reach any conclusions, the following questions had to be answered:

1. How reliable, or consistent were the diagnoses?
2. How reliable were the remediations?
3. How reliable were information- (cue-) collection procedures?
4. How reliable were diagnostic statements to which remediations had been attached?
5. How reliable were specific remedial/diagnostic associations?
6. How strong was the correlation between diagnostic and remedial performance?
7. Was there a discernable institutional training effect on performance?
8. How did the specialists themselves view the tasks of arriving at a diagnosis and initial remediation plan?

To obtain the necessary data, eight experienced, credentialled, practicing reading specialists, four from Michigan and four from Illinois, were asked to examine three simulated cases of reading difficulty in order to (1) make diagnostic judgments, (2) prepare initial remediation plans, and (3) associate given remediations with diagnostic statements. In all, 24 sessions were conducted.

Design and Method

The subjects, all women, were randomly assigned to simulated cases of reading difficulty. Each case was based on a real child who had attended the Michigan State University Reading Clinic. The four simulated cases developed by the Clinical Studies Project were considered to be representative of reading problems commonly encountered in public schools. The children's grade levels in the cases ranged from third to seventh. In the judgment of a senior reading clinician on the faculty of Michigan State University, the cases exhibited the following problems:

> Case One: (1) significantly depressed sight vocabulary, (2) serious problem with decoded word recognition skills, (3) inadequate fluent message segmentation, and (4) no significant problem comprehending anything read or listened to.
>
> Case Two: (1) significantly retarded sight vocabulary and word-recognition behaviors, (2) significant problem with analysis on multisyllabic words coupled with poor visual segmentation and high-frequency hearing loss, (3) difficulty phrasing material of high level semantic content, and (4) comprehension problems related to the demands of higher level content-related materials.
>
> Case Three: (1) reasonably intact sight vocabulary and word recognition performance, (2) significant problems with higher level decoding skills and their application during reading, (3) serious problem with fluent texting of material, and (4) comprehension problem specifically related to the meaning demands of content-related materials.
>
> Case Four: (1) significantly depressed sight vocabulary, (2) significant problem with the learning and application of decoded word recognition skills, (3) no major problem with text segmenting, and (4) no language-based comprehension problem.

For each simulated case a cue inventory was provided listing the information (cues) available for that case: achievement tests, family and academic background, cognitive ability, group and individual reading diagnostic measures, classroom information, work samples, and so on. The information was presented in a variety of formats: test booklets, audio tapes, examiner's comments, and test scores. All cases contained an audio-taped interview with the student, a brief statement of the reason for referral to the Reading Clinic, and an artist's sketch of the child based on the

artist's impression of the child after hearing a taped interview of him/her.

Each simulated case had an equivalent form--a superficially disguised replicate of the original prepared by making minor changes in the data base and randomly re-ordering the cue inventory (Lee & Weinshank, Note 1). Thus there were eight cases, four originals and four replicates.

The subjects' task, in all the sessions, was to diagnose the simulated case at hand by requesting information from the cue inventory, propose an initial remediation plan, and associate remedial and diagnostic statements. Specifically, the procedures were as follows:

1. The clinicians coded each diagnostic statement as being a strength, weakness, or observation and transferred the statements to a standardized diagnostic checklist.

2. They coded each remedial statement as a strength, weakness, observation, or treatment, and transferred the statements to a standardized remedial checklist.

3. They associated remedial and diagnostic statements.

4. They explained what guided them in gathering information about a child and in determining what should be written in the diagnosis. Further, they discussed what they believed ought to be the relationship between diagnosis and remediation.

Table 1 summarizes the content of each session. A minimum of one week separated sessions one and three.

Table 1.
Session Content

Session 1	Session 2	Session 3
Work with one of the four original simulated cases.	Work with one of the remaining three original simulated cases.	Work with the equivalent form (replicate) of the Session 1 simulated case.

Data Analysis

Data from the 24 observational sessions (three sessions for each of the eight subjects) were analyzed on a case-by-case basis (three clinicians/

six sessions per case). Agreement between and among clinicians on statements characterizing each case was measured using a proportional agreement statistic for five products: diagnostic statements, remedial statements, cues collected, diagnostic statements seen as requiring remediation, and specific diagnostic/remedial associations. Agreement (the proportion of clinicians agreeing on statements seen as characterizing each case) was measured and Phi Correlations (the extent of agreement between a clinician and herself over time, and between two clinicians on statements believed to characterize a case) were computed. Additionally, strength of association between cues and diagnosis, cues and remediation, diagnosis and remediation, and so on was measured using Pearson Product-Moment correlations. Difference in performance between clinicians trained in Michigan and Illinois was measured using t-tests of differences between mean performances of both groups.

Results

Most of the statements made about a case were made only once; they were idiosyncratic. Setting the most generous standard for reliability--that the statement be mentioned in two or more of the six sessions devoted to each case--resulted in clinician agreement on only one-third or fewer of the total statements made (excluding cues). When the standard for reliability was changed to reflect those statements made in three or more of the sessions, agreement plunged virtually to chance level (see Figure 1). (The horizontal axis in Figure 1 shows the six sessions as proportions, (e.g., 0.33 means two sessions).

Group Agreement

Diagnosis. Across the four cases, a total of 304 diagnostic statements was made. Thirty-three of these statements, clustered within nine categories, accounted for all statements made in two or more sessions.

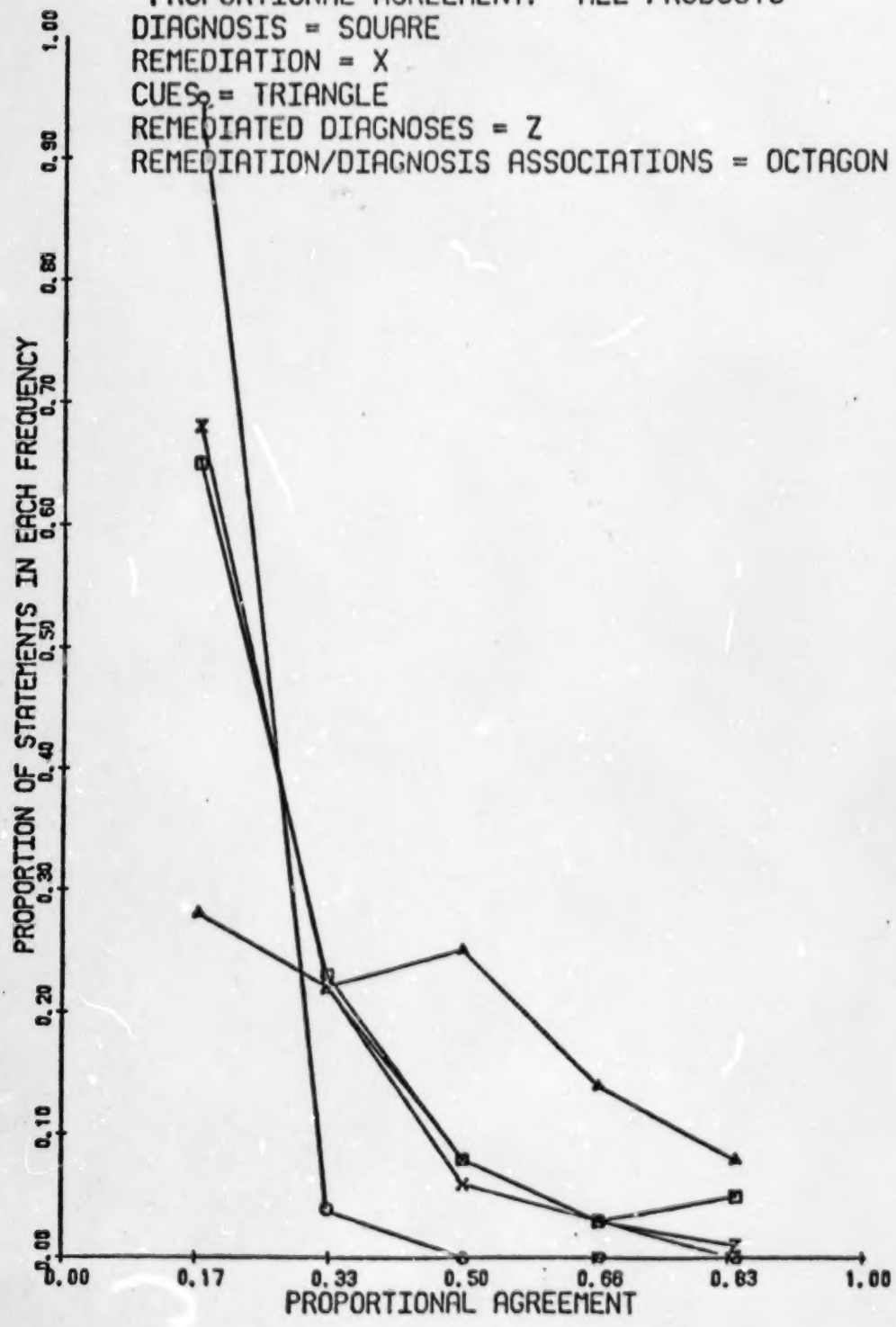

Figure 1. Proportional agreement.

The nine categories are intellectual potential, word recognition, word analysis, oral reading, silent reading, comprehension, hearing, vision, and affect. What group agreement did exist, then, centered entirely around these diagnostic categories.

The two sets of categories--those mentioned in two or more sessions and those mentioned only once--had different characteristics. The proportion of statements coded as weaknesses was substantially higher for the categories mentioned in two or more sessions than for those mentioned only once. The idiosyncratic statements, that made up the bulk of the diagnostic statements for all the cases, tended to revolve more around statements of strengths and observations. Those statements made most often across sessions tended to focus more sharply on problematic aspects of the child's reading performance (see Table 2).

Remediation. A total of 270 remedial statements were made across the four cases. Twenty-four of these statements, clustered within seven categories, accounted for all statements made in two or more sessions. The seven categories are sight words, phonetic analysis, structural analysis, oral reading, visual, comprehension, and motivation (see Table 3).

Again, as with the diagnoses, the two sets of categories--those mentioned in two or more sessions and those mentioned only once--had different characteristics. There were more treatment statements in the set of remedial statements made in two or more sessions than in the set of remedial statements made only once. Many of the idiosyncratic statements were essentially diagnostic rather than remedial. That is, they were coded as strengths, weaknesses, or observations rather than as treatments. Those statements made in two or more sessions focused almost exclusively on treatment actions.

Table 2.

Agreed-Upon Diagnostic Statements Across Cases

Cases				Statements
1	2	3	4	
X			X	Sight words: identify high frequency (context)
X	X		X	build sight vacabulary
X				general
X			X	use in language experience
			X	use sight words/phrases
X				Oral reading: improve rate
X				general
X		X		improve fluency
	X		X	Phonetic analysis: general
X				apply phonics to attack new words
X			X	drill on short/long vowel generalizations
			X	blending, general
			X	blend sound units
X				Structural analysis: general
X			X	recognize common word parts
	X			add prefixes/suffixes
	X			count syllables
	X			identify root words
	X			teach syllables
			X	Vision: discriminate visually similar words
	X	X	X	Comprehension: general
	X	X		information gathering
		X		demonstrate factual recall
X	X	X		use context as aid in comprehension
			X	Motivation: general
		X		setting own purpose
	X		X	provide high interest materials

Table 3.

Agreed-Upon Remedial Statements Across Cases

Cases				Statements
1	2	3	4	
X	X	X	X	Intellectual potential
X			X	Word recognition: general
X			X	sight
X				reversals
X	X			Oral reading: general
X			X	rate
X				punctuation
X		X		substitutions contextually acceptable
X		X		phrasing
X				fluency
X	X		X	Word analysis: general
	X			suffixes
	X		X	letter/sound associations
			X	initial consonants
X		X	X	blending ability
	X	X		syllables
		X		use of blends
		X		use of root words
	X			Hearing: discrimination
	X			general
	X	X		acuity
		X	X	Vision: acuity
X	X	X	X	Silent reading: rate
	X	X		comprehension
		X	X	Comprehension: oral
		X	X	general
X	X	X	X	silent
		X	X	listening
X				Affective: general
		X	X	attitude toward reading
			X	home environment
		X		relationship with peers
		X		adjustment

Cues. On the average, there was considerably more agreement on which cues ought to be collected than on which diagnostic and remedial statements characterized a case. As with the diagnoses and remediations, however, the two sets of cues had different characteristics. Those cues requested in two or more sessions presented performance information directly (as test scores and as examiner's comments). Cues mentioned in only one session presented information in the relatively high inference form of test booklets and audio tapes, which required that the subjects make judgments about the child's problems based on an examination of the child's performance on the protocols.

The test scores and comments requested in two or more sessions dealt with background (home/school), cognitive ability, diagnostic tests of reading ability, graded word lists, basic sight vocabulary, reading achievement tests, auditory discrimination tests, and reports of visual/auditory acuity. Some clinicians collected many items of information within each category, others collected very few items. Thoroughness of collection had little effect on diagnostic reliability, however. The clinician who collected the fewest items of information ($\bar{X}$=15) had a diagnostic reliability of 0.00. The clinician who collected the most items ($\bar{X}$=44) had a diagnostic reliability of only 0.12.

While categories of cues were collected consistently across cases, agreement within cases on specific diagnostic and remedial statements was extremely limited. Across cases, then, common cues resulted in largely non-common diagnoses and non-common remediations, indicating a lack of mutually agreed-upon interpretations of common information sources.

Remediated diagnoses. About two-thirds of all diagnostic statements were seen as requiring remediation, but few were repeated two or more times.

In other words, the clinicians identified diagnostic statements that they felt required remedial attention but did not agree on a common set of such diagnostic statements for each case. As was reported earlier, diagnostic statements coded as weaknesses were more likely to be remediated than those coded as either strengths or observations.

Specific remedial/diagnostic associations. Across the four cases, only 31 specific remedial/diagnostic associations out of the 654 made were repeated across two or more sessions. Put another way, from 94% to 98% of all specific remedial/diagnostic associations were mentioned in only one session. The few associations that were made in two or more sessions tended to pair treatment with weakness.

Intercorrelations

The likelihood of any two clinicians agreeing on statements characterizing a given case (inter-clinician correlation) ranged from small to non-existent. The likelihood of clinicians agreeing with themselves about a case and its replicate (intra-clinician correlation) was only slightly more favorable than for two separate clinicians. A partial intra-clinician correlation matrix is presented in Table 4. Only four statements out of 30 made were agreed upon by the same clinician for the same case at two different points in time.

The mean inter- and intra-clinician correlations are summarized in Table 5. Table 5 illustrates the very limited agreement within and between clinicians on the cases. Even at its best, clinician reliability was markedly inadequate.

Table 4.

Partial Intracorrelation Matrix for Diagnosis
Case 3, Clinician A, Time One,
Clinician A, Time Two

STATEMENT ++

Present at Time One and Time Two

Motor coordination (W)
Intellectual potential: General (S)
Oral reading: Phrasing (W)
Oral reading: Intonation (W)

STATEMENT +-

Present at Time One, Absent at Time Two

Attitude toward reading: Independent (obs)
Motivation for reading (Obs)
Emotional adjustment (W)
Substitutions contextually acceptable (W)
Silent reading comprehension (W)
Word analysis (W)
Phonetic analysis (W)
Comprehension vocabulary (Obs)

STATEMENT -+

Absent at Time One, Present at Time Two

Motor coordination (Obs)
Hearing acuity (W)
Speech articulation (W, Obs)
Attitude toward reading: Independent (W)
Attitude toward reading: Instructional (W)
Relationship to peers (W)
Ability to apply reading skills (W)
Oral reading: General (Obs)
Oral reading: Rate (W)
Oral reading: Self-correction (W)
Silent reading: General (W)
Word analysis: General (S)
Phonetic analysis: General (Obs)
Use of initial consonant sounds (S)
Use of syllables (S)
Word recognition: General (S)
Comprehension: Oral (S)
Comprehension: Listening (S)

STATEMENT --

Absent Both at Time One and Time Two

272 Domain statements excluded

Table 5.
Mean Inter-Clinician Correlations Across Cases

Cases 1-4	Inter-Clinician Correlation	Intra-Clinician Correlation
Diagnosis	0.11	0.14
Remediation	0.10	0.20
Cues	0.29	0.49
Remediated Diagnoses	0.08	0.14
Remedial/Diagnostic Associations	0.00	0.02

Strength of Association (Correlational Data): Individual and Group

At the individual clinician level, performance on one product could not be used to predict performance on another. Cues were predictors neither for the diagnosis ($r = 0.27$, $r^2 = 0.07$), nor for the remediation ($r = 0.25$, $r^2 = 0.06$). Neither was the diagnosis a predictor for the remediation ($r = 0.10$, $r^2 = 0.01$).

At the group level, only diagnosis and remediation showed modest levels of association ($r = -0.58$, $r^2 = 0.33$). Cues did not predict either diagnosis ($r = 0.14$, $r^2 = 0.02$) or remediation ($r = -0.18$, $r^2 = 0.03$), even at this level.

Comparison of Subjects

Five t-tests for differences between means were computed to test the hypothesis that there was no difference between the performance of the Michigan and the Illinois subjects. None of the tests reached statistical significance at the .05 level, but the difference in cues did approach significance. The Michigan subjects requested fewer cues (231 vs. 419) and, proportionally, more high-inference cue forms (0.42 vs. 0.09) than did the Illinois subjects. Nevertheless, the differences did not translate into different diagnostic or remedial performance. Statistically,

the groups were virtually identical in diagnostic and remedial performance, indicating that irrespective of number or form of cues collected, there was no common agreed-upon heuristic to guide information gathering and interpretation.

Clinician Self-Reports

Each clinician reported on plans of action that she used to guide her information gathering, the writing of her diagnosis, and her initial remediation. The clinicians reported collecting many of the same types of information (e.g., on intelligence scores, comprehension, word attack and recognition skills). Further, they shared very similar procedures for writing their diagnosis. They said that they organized their diagnosis either (1) in the order in which they collected information from the children, (2) by grouping their findings by strengths and weaknesses, or (3) by listing findings according to areas of need, together with supporting information. Finally, they held similar views as to the relationship that ought to exist between diagnosis and remediation. They asserted that (1) There should be a direct one-to-one relationship between the two, although there will be some things that cannot be remediated; (2) Weaknesses should be treated by using strengths; (3) Weaknesses should be prescribed for directly.

The subjects' responses to all three questions were prompt and very similar, suggesting that they had internalized a standard rhetoric for describing the task of diagnosis and remediation.

An examination of teacher performance on cue collection, diagnostic write-up, and remedial write-up, however, revealed that actual performance was never consonant with the stated plan.

Across the eight clinicians, two overall observations were made regarding cue collection.

1. The order in which cues were actually collected across sessions did not follow the order given in the stated plan. Further, whatever order was used, it was never followed a second time, either for the middle (distractor) case or the replicate case.

2. A substantial amount of information was collected that was never mentioned in the stated plan at all. Six of the the eight clinicians requested this additional information to the extent that it equalled or exceeded the information that was collected according to plan.

While it was possible that this information might have been requested merely because it was available, the data warrant the interpretation that the clinicians had no ready framework within which to fit the new data, and had no commonly shared mode of interpreting what they did request.

Across the eight clinicians, three overall observations were made regarding diagnostic write-ups.

1. In no instance did the diagnostic statements follow the order of cues collected.

2. Whatever the order in which diagnostic statements were presented, no clinician repeated that order a second time.

3. In those cases where strengths and weaknesses were loosely grouped and supporting information was provided, the pattern was never the same for all three sessions for a clinician, and the order of presentation always varied.

Finally, three overall observations were made regarding the associations made between remediation and diagnosis.

1. In no instance did the remedial statements follow the order of presentation of the diagnosis.

2. There was no evidence of systematic use of weaknesses being treated by using strengths.

3. No clinician had a one-to-one match between diagnosis and remediation; there were always unmatched diagnostic statements.

Discussion

The assumption in the literature that diagnosis and remediation are, or ought to be, routinely and reliably associated is not remotely substantiated by any of the findings of this study except at the case level. Even at that level, the association is only modest.

Across simulated cases, clinicians examined information about student home and school background and cognitive ability, and the results of performance on diagnostic tests, grade word lists, reading-achievement tests, basic sight vocabulary, and visual/auditory tests. Thus they all began with essentially the same categories of information for all the cases, although the order of collection was entirely unpredictable.

As the data reported has shown, there was little association between diagnosis and remediation. A glaring example of this is Case 3, as shown in Tables 2 and 3. Seventeen diagnostic statements were agreed upon as characterizing the case. Only six remediations were agreed upon as constituting an appropriate action plan, and of those, two were unrelated to any of the agreed-upon diagnoses. Additionally, with the exception of Case 1, the group diagnoses were at variance with the model diagnoses compiled by the senior reading clinician. Differences included omission of weakness in fluent texting and comprehension for Case 2; omission of weakness in sight vocabulary and phonetic analysis for Case 3; and omission of weakness in multisyllabic analysis, fluent texting, and auditory problems for Case 4.

Thus the clinicians may have agreed as a group that certain problems characterized a case,but, as a group, they may have been largely incorrect in their judgments. Conversely, they may have compiled correct group

diagnoses and the judgments of the senior clinician may have been in error. The question of the validity of the diagnoses (as distinguished from their reliability--one can be reliably wrong) was beyond the scope of this investigation. However, it is clearly a major unresolved issue.

The data in this study unambigously demonstrated that the clinicians studied were not reliable diagnosticians. They believed they were acting consistently, but in fact they collected a great deal of information in an unpredictable fashion, and interpreted most of it idiosyncratically.

Why did the clinicians in this study behave as they did?

Speculation 1: Task and External Validity Contributed to Low Reliability

The subjects' performance might have been an artifact of the experimental setting, including the use of simulated cases. That is, the task itself was crafted so that while the clinicians could perform comfortably as required, they weren't doing what they actually do on the job. Perhaps a natural setting using real children would have yielded different results. However, any unreliability demonstrated in a natural setting could be attributed to the differing behaviors of the children themselves. Both settings, then, contribute to problems of task validity. Nevertheless, a major assumption in this study, as well as in training and research studies in medicine, is that important problem-solving behaviors of clinicians can be elicited through the use of simulated cases. "At least some of the important cognitive skills are elicited by sets of data gathered about real cases with real problems and stored for presentation as simulated cases" (Vinsonhaler, Note 2).

The question of external validity, or generalizability to a larger population, is complicated by the small sample size. However, a major trade-off for small sample size is the power of replication. The clinical studies group has conducted a number of studies of experienced

reading specialists. The conditions were similar to those used in the present study and the diagnostic results were almost identical to those reported here (Vinsonhaler, Note 2; Gil, Note 3; Hoffmeyer, Note 4; Stratoudakis, Note 5). While this is the first study to investigate remedial reliability and the reliability of remedial/diagnostic associations, there is no reason, in principle, to think that these results would not be replicated with another sample.

A second issue in external validity is the nature of the sample itself. Was this a representative sample of experienced, credentialled reading specialists, or did chance dictate selection of an unusually incompetent and unrepresentative group? These subjects met or exceeded the criteria set by the study. They were all employed in good standing in five different school districts in two states. They were involved in professional reading organizations and/or had been recommended by their academic mentors as having been highly competent students. If anything, one might argue that they were an exceptionally able group, thus making the results all the more disquieting.

Speculation 2: Training Influenced Performance

Contemporary teacher training programs are committed to a position that calls for attending to individual differences ("meeting individual needs"). This view is not compatible with the proposition that a small number of diagnostic and remedial categories account for a large proportion of the problems exhibited by deficient readers of a given age and grade-level. Training that embodies the individual differences paradigm has no need of predictive capability. By definition, if everyone is different in significant ways that must be uncovered, one student's performance cannot and should not have any relationship to another student's performance. If each student's reading problem is approached from the

standpoint that the problem is different in significant ways from everyone else's, those differences must somehow be divined from the cues. Which cues are the key ones? There is no evidence in this study that the training process has addressed itself to this question, as shown by the almost complete lack of diagnostic and remedial agreement among clinicians and between each clinician with herself on the same case.

If the individual-differences speculation is correct, then the clinicians in this study are performing precisely as they have been trained. Making between 20 and 30 diagnostic statements and between 25 and 35 remedial statements per session (as almost all did) would surely lead a clinician to believe that individual needs had been taken fully into consideration.

Speculation 3: The Clinical Setting Itself Does Not Provide a Context for the Specialist to Learn from Experience

Diagnostic judgments are reached and remedial suggestions made, but there is no way for the clinician to get feedback about the consequences of recommended actions. The clinician simply does what the profession dictates with no opportunity to confirm or disconfirm hypotheses about problems and their treatment.

Speculation 4: It Is the Nature of the Profession to Mystify the Process so that it Merits the Respect it Claims for Itself

Perhaps the painful reality is that most diagnoses are embarassingly inflated and clinicians know that such diagnoses ought not to be taken seriously when it comes to remediation. Similarly, clinicians may understand all too well the limited range of action alternatives open to them, but may feel that professional mystique compels them to engage in clinical overkill.

Speculation 5: The Clinicians Had a Plan But They Didn't Have a Satisfactory Method

According to Newell and Simon, (1972), "A method is simply a plan that you use twice" (p. 835). The plans the clinicians believed they were following to organize their behavior were never actually carried out twice in the same way; there was no way to predict clinician behavior from case to case. The method actually used may be the answer to the question, "How can you make a straight-forward instructional problem look very complicated? Answer: disguise the fact that, in general, a very limited number of diagnostic and remedial judgments reliably characterize most cases of reading difficulty. Produce a verbal jungle of strengths, observations, and weaknesses derived from a needlessly exhaustive examination of case information collected unsystematically. This procedure is guaranteed to take quite a bit of time, look substantive, and impart a feeling of professionalism.

Speculation 6: The Clinicians Have Not Mapped Onto the Task Any Model of the Process of Acquisition and Mastery of Reading Behaviors.

Lacking such a model, their search for, and interpretation of, information is unordered, and their resulting instructional strategies are unsequenced.

None of these speculations is flattering to the field of remedial reading. But the results of this study are such that unless one were to assert that the subjects were "lemons" and the experimental conditions spurious (charges that have been examined and refuted), the near-chaotic state of diagnosis and remediation cannot go unchallenged.

Conclusions

As currently practiced, the compilation and presentation of diagnostic findings seems to be a random exercise. There are as many different conclusions for a case as there are points of entry into the data. The bulk of the diagnostic statements are idiosyncratic, and hence unpredictable. The bulk of the remedial statements are idiosyncratic, and may or may not (as chance dictates) have an impact on a child's reading performance. Given the very low level of diagnostic and remedial predictability demonstrated in this study, it can be argued that diagnostically derived remediation, as presently conducted, is neither time efficient, economically justifiable, nor educationally credible.

Reference Notes

1. Lee, A., & Weinshank, A. Case production and analysis: CLIPIR pilot observational study of reading diagnosticians (Research Series No. 14). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1978.

2. Vinsonhaler, J. The consistency of reading diagnosis (Research Series No. 28). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

3. Gil, D. The clinical problem solving behavior of classroom teachers as they diagnose a child's reading performance in laboratory and classroom. Unpublished doctoral dissertation, Michigan State University, 1979.

4. Hoffmeyer, L. The process and outcomes of diagnostic problem solving among reading clinicians. Unpublished doctoral dissertation, Michigan State University, 1980.

5. Stratoudakis, J. The construction and use of paper problems to observe the diagnostic problem solving behavior of reading clinicians. Unpublished doctoral dissertation, Michigan State University, 1980.

References

Newell, A., & Simon, H. Human problem solving. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1972.